

# MISO 0.1.6RC User Manual

- 1 [MISO Overview](#)
  - 1.1 [Introduction](#)
  - 1.2 [Overview of Features](#)
  - 1.3 [Implementation](#)
  - 1.4 [Releases](#)
  - 1.5 [Licensing](#)
- 2 [Getting Started](#)
  - 2.1 [Installation and running MISO](#)
  - 2.2 [Logging into MISO](#)
    - 2.2.1 [External users](#)
    - 2.2.2 [Internal users](#)
  - 2.3 [Dashboard](#)
    - 2.3.1 [Home](#)
    - 2.3.2 [My Account](#)
    - 2.3.3 [My Projects](#)
    - 2.3.4 [Reports](#)
    - 2.3.5 [Analysis](#)
    - 2.3.6 [Help](#)
- 3 [Configuring MISO](#)
  - 3.1 [Users and Groups](#)
    - 3.1.1 [Listing Users](#)
    - 3.1.2 [Creating Users](#)
    - 3.1.3 [Listing Groups](#)
    - 3.1.4 [Creating Groups](#)
  - 3.2 [MISO Configuration](#)
    - 3.2.1 [General](#)
    - 3.2.2 [Databases](#)
    - 3.2.3 [Security](#)
    - 3.2.4 [Barcode Printers](#)
  - 3.3 [Alerts](#)
    - 3.3.1 [My Alerts](#)
      - 3.3.1.1 [Project Alerts](#)
      - 3.3.1.2 [Run Alerts](#)
      - 3.3.1.3 [Pool Alerts](#)
    - 3.3.2 [Recent Activity Alerts](#)
  - 3.4 [Cache Administration](#)
    - 3.4.1 [Flush All Caches](#)
    - 3.4.2 [View Cache Stats](#)
    - 3.4.3 [Regenerate Barcodes](#)
    - 3.4.4 [Reindex Alert Managers](#)
  - 3.5 [Configuring Sequencer References](#)
    - 3.5.1 [Viewing a Sequencer Reference](#)
    - 3.5.2 [Adding a Sequencer Reference](#)
    - 3.5.3 [Editing a Sequencer Reference](#)
- 4 [Working with MISO](#)
  - 4.1 [Project](#)
    - 4.1.1 [Creating a project](#)
    - 4.1.2 [Adding a Project Overview](#)
    - 4.1.3 [Samples](#)
    - 4.1.4 [Libraries](#)
    - 4.1.5 [Library Dilutions](#)
    - 4.1.6 [EmPCRs](#)
    - 4.1.7 [EmPCR Dilutions](#)
    - 4.1.8 [Pools](#)
    - 4.1.9 [Studies](#)
    - 4.1.10 [Runs](#)
  - 4.2 [Sample](#)
    - 4.2.1 [Creating a single sample](#)
    - 4.2.2 [Bulk sample creation](#)
    - 4.2.3 [Sample operations](#)
      - 4.2.3.1 [Printing Sample Barcodes](#)
      - 4.2.3.2 [Adding Notes](#)
      - 4.2.3.3 [Adding Sample QCs](#)
      - 4.2.3.4 [Adding Libraries](#)
    - 4.2.4 [Receiving samples](#)
  - 4.3 [Library](#)
    - 4.3.1 [Adding Libraries](#)
    - 4.3.2 [Bulk Library creation](#)
    - 4.3.3 [Library Operations](#)
      - 4.3.3.1 [Adding Notes](#)

- 4.3.3.2 Adding Library QCs
    - 4.3.3.3 Add Library Dilutions
    - 4.3.3.4 Constructing Pools
  - 4.4 Pool
    - 4.4.1 Creating Pools
    - 4.4.2 Pool Operations
  - 4.5 Study
    - 4.5.1 Creating a study
    - 4.5.2 Edit Study
      - 4.5.2.1 Adding Experiments
  - 4.6 Experiment
    - 4.6.1 Automatic Experiment Creation
    - 4.6.2 Add Experiment Manually
  - 4.7 Partition Containers
    - 4.7.1 Adding Partition Containers (outside a Run)
    - 4.7.2 Adding Partition Containers (within a Run)
  - 4.8 Run
    - 4.8.1 Automatic Run Notification
    - 4.8.2 Adding a Run Manually
    - 4.8.3 Run Operations
      - 4.8.3.1 Editing Partition Containers
      - 4.8.3.2 Adding Notes
      - 4.8.3.3 Adding Run QCs
- 5 Reporting
  - 5.1.1 Generating Reports
  - 5.1.2 Report output
  - 5.2 Sequencer Resource Calendar
- 6 Submissions
  - 6.1 Preparing new Submissions
  - 6.2 Previewing, Validating and Submitting Metadata
  - 6.3 Submitting Sequence Data
  - 6.4 List Submissions
- 7 Notification Server
  - 7.1 Installing
  - 7.2 Configuring
  - 7.3 Running
- 8 Analysis Server
  - 8.1 Installing
  - 8.2 Configuring
  - 8.3 Running
- 9 Troubleshooting

## MISO Overview

### Introduction

MISO (Managing Information for Sequencing Operations) is the Lab Information Management System developed at TGAC. Work started on MISO in 2010 and was released to the community as an open-source LIMS on the 31st May 2012. MISO is heavily modelled upon the submission schemas specified by the SRA and is therefore able to automatically generate and deploy the required XML and data files.

### Overview of Features

- **Authentication** – user-centric access control to designated areas Laboratory tracking – project description, sample receipt, library preparation, run construction, and barcoding
- **Printing** – can directly connect to barcode printers and print barcodes for the relevant objects inside MISO
- **Bioinformatics pipelines** – monitoring and reporting of analytical processes, interacting directly with the computing cluster
- **Reporting** – personalised and filtered accurate statistics from library preparation and sequencing processes and including graphical visualisations
- **Data Visualisation** – through “traffic light” indicators, tree-structured status diagrams and statistics graphs and plots, MISO can easily show the status of each project and its elements by interactive tree and dendrogram, as well as calendar-based sequencer resource diagrams and run statistics, all using the JavaScript graphing library d3.js.
- **Submission** – automated packaging of sequencing data and metadata, deployment of these data to remote public repositories (i.e. EBI)
- **Notification** – automated import of run metadata and notification of change of run status

## Implementation

Full model and implementation documentation can be found in the [Developer Manual](#). Some brief MISO facts:

- Currently MySQL underpins MISO. Object relational mapping (ORM) via data access objects (DAOs) is achieved through Spring's JDBC Templates, which enable fine-grained control over queries over the underlying database.
- Simple Java beans, or Plain Old Java Objects (POJOs) are used to represent the model objects in the database, and we construct concrete relationships between them in business logic space.
- We have developed a RESTful web application from the Spring 3 MVC framework and, as such, MISO can support a number of view technologies. It comes with a user-friendly interface built on Java Server Pages (JSP) which is designed specifically for lab technicians, but can be queried programmatically using the REST URLs. This means helpful web services, such as remote custom reporting and integration with other tools and instrumentation, are available.
- [D3.js](#) is used for much of the graphical data visualisation in MISO.

## Releases

- MISO has been developed by the Core Bioinformatics Team at TGAC since February 2010 under leadership of Dr. Robert Davey.
- The first version of MISO (version 0.1.1, Hydrogen) was released internally to TGAC staff on 19th Sept 2011. Each month brings a new version with novel features and bug fixes. The latest internal release was on 30-06-2012 Version: 0.1.6 (Carbon), and this release represents what will be the first beta community release of MISO.

## Licensing

- MISO is open-source beta software at present, and is freely available for usage and modification according to the GNU Public Licence V3 (GPL3).
- We use freely available open-source tools to develop MISO and, in turn, MISO is fully open-source.
- The Spring framework enables us to supply configurations via XML files, so functionality changes are possible without any code changes. For example, replacing a MySQL database model with Postgres or replacing the authentication mechanism is possible with minimal configuration.
- Coupled with the open-source ethos of community development, MISO can make for a cost-effective, highly modifiable and robust system that can benefit both the public knowledge repositories and sequencing centres of all sizes.
- We are actively working on MISO to provide more features, and to support more sequencing platforms.

## Getting Started

### Installation and running MISO

- The easiest way to install MISO is to install the free [VirtualBox](#) software and grab our specially crafted virtual machine appliance from: [https://repos.tgac.bbsrc.ac.uk/miso/MISO\\_0.1.6RC\\_Image.ova](https://repos.tgac.bbsrc.ac.uk/miso/MISO_0.1.6RC_Image.ova)
- Import this OVA file into VirtualBox via the *File, Import Appliance* option.
- Starting the appliance will result in an Ubuntu virtual machine complete with all MISO's prerequisites pre-installed. The login for the virtual image is **miso / misoadmin**
- Log in to MISO using **admin / admin**. Only some simple configuration for certain elements is required, such as Sequencer References, Printers and any additional users and groups.
- As MISO is open-source, the complete source code is available from
  - github: <https://www.github.com/TGAC/miso-lims>
- The build, install and running instructions can be found in the [Developer Manual](#).

### Logging into MISO

MISO ships with a default admin account after installation, using the username password combination of **admin/admin**. To add more users and groups please see the section below. For information regarding running MISO using the database or LDAP authentication, please see the Developer Manual.

#### External users

- External users have limited access to MISO, via a restricted read-only interface, which shows the list of projects with which they are involved
- This interface comprises detailed information about which stage the project is currently at, and QC information about Samples, Libraries and Runs.

#### Internal users

- Internal users can access most of the features of MISO, depending on the roles allocated to them via the internal user and group management system, or another user directory such as LDAP.

## Dashboard

### Home

- The Home area represents the MISO dashboard, and includes widget lists of Projects, Runs, Samples, Library, etc, that the logged-in user is able to read.
- The input fields on the top of each widget can be used to search the relevant object list based on name, alias, description and barcode.

### My Account

- The My Account tab lets the user access areas of MISO via the sidebar navigation menu, or via the widgets in the main section of the screen.
- This tab also includes the user's account settings and configuration. Administrators can see the full [configuration and administration](#) and configuration settings of MISO.

### My Projects

- This tab lists all the projects available in MISO that are viewable by the logged-in user, with filtering functionality and provides two diagrams showing all projects and objects they contain in a round tree or dendogram style.

### Reports

- Reporting is an important functionality of MISO, with different type of reports being able to be generated based on Projects, Runs, Samples, Libraries, and the resource calendar of sequencing machines.

### Analysis

- The analysis area provides access to the MISO analysis pipeline system, which allows custom pipelines to be started on a computing cluster via a web interface.

### Help

- [Link to the user manual.](#)

## Configuring MISO

## Users and Groups

### Listing Users

- Click the *List Users* link in the Navigation Panel. You will see a list of currently registered Users.
- Click the *Edit* link to start modifying a User.
- Click the *Add User* link to add a new User.

### Creating Users

- Add the required fields in the form for the user.
- A user can either be **Internal** or **External**
  - An Internal user will have access to the full MISO application, depending on their permissions.
  - An External user will have access to a special user interface of MISO that only lets that user view and complete specific elements, depending on their permissions.
- Users can easily be added to Groups by simply checking the boxes next to their names in the Groups checklist.
- Click *Save* to save the User.

### Listing Groups

- Click the *List Groups* link in the Navigation Panel. You will see a list of currently available Groups.
- Click the *Edit* link to start modifying a Group.

- Click the *Add Group* link to add a new Group
- Three groups should exist in MISO already, and they are **ProjectWatchers**, **RunWatchers**, and **PoolWatchers**. These groups represent the ability to watch all elements of a given type for alerting updates, and as such shouldn't be removed.

## Creating Groups

- Add a Name and Description for the new Group.
- Users can easily be added to Groups by simply checking the boxes next to their names in the Users checklist.
- Click *Save* to save the Group.

## MISO Configuration

Some elements of MISO can be modified through the MISO Configuration widget in the My Account page of administrators.

### General

Unused at present. This area will be populated with runtime modifiable options.

### Databases

Unused at present. This area will be populated with runtime modifiable options. Databases can be configured through the JNDI which is detailed in the Installation Manual.

### Security

Unused at present. This area will be populated with runtime modifiable options. Security is based on user and user group, with defined permissions for each object in the MISO.

### Barcode Printers

- Barcode printers can be added and managed from this page.
- Clicking the *Add Printer Service* link will add a row to the table with relevant information for the printer to add.
- Type in a unique **Printer Service Name**, select a printer context **type** and options will appear for that type, then choose the element that this printer will print.
- To print multiple element types from the same printer, simply add separate services for each of the element types, pointing to the same host.

## Alerts

### My Alerts

MISO has an alerting system whereby users can watch singular elements or whole groups of elements and receive alerts about them. Depending on the way MISO is configured, you will either receive internal MISO alerts, accessed through the My Account tab, via email, or both.

### Project Alerts

- To receive Project level alerts, you will need to be one of the following:
  - watching a singular Project by selecting *Watch* in the Option menu of a Project
  - the owner of a Project
  - in the ProjectWatchers group

### Run Alerts

- To receive Run level alerts, you will need to be one of the following:
  - watching a singular Run by selecting *Watch* in the Option menu of a Project
  - the owner of a Run
  - in the RunWatchers group

### Pool Alerts

- To receive Pool level alerts, you will need to be one of the following:
  - be the owner of a Pool
  - in the PoolWatchers group

## Recent Activity Alerts

If you are and Administrator, you will be able to see the list of recently generated system level alerts. This represents a traceable timeline of events for given elements, and future versions of MISO will be able to utilise these events to build up graphical representations of project flow.

## Cache Administration

### Flush All Caches

- Flushing all caches can be use to clear all the locally stored POJO objects in memory. This is useful if an administrator has changed a reference value (i.e. a Sample alias) manually in the underlying database where the original value would still be held in the in-memory cache.

### View Cache Stats

- The statistics of each of the caches can be viewed by clicking this link.

### Regenerate Barcodes

- Generates all the barcodes again in the storage area on disk from the barcode strings of each relevant barcodable object.

### Reindex Alert Managers

- The MISO alerting system keeps a copy of objects watched by the alerting system. This also helps MISO in being able to form an index of these objects for fast searching. This link will reindex all objects held by this system.

## Configuring Sequencer References

Sequencer References represent the link between a physical sequencer and MISO, and are essential for the input of Runs into the system.

### Viewing a Sequencer Reference

- The All Sequencer References page shows the currently available sequencers linked to MISO.
- By clicking on the *View* link of a sequencer reference, all the runs performed on that machine can be seen with corresponding name, health, date and status.

### Adding a Sequencer Reference

- When a new sequencer reference is to be added, click on the *Add Sequencer Reference* link in the top-right of the All Sequencer References page.
- A new row will be added to the table, where an ID for the instrument will be generated automatically.
- A unique name **must** be supplied for the sequencer reference, and it **must be** the manufacturer specified name of this instrument, e.g. for an Illumina run folder 120313\_SN319\_0209\_AD0HPAACXX, the sequencer reference name would be SN319. If the name is incorrect here, the notification system will not be able to automatically add runs to MISO.
- Select a platform for the instrument
- Set the hostname of the machine. This will either be the head node PC of a given sequencer, or more likely the hostname of a cluster or compute area where the run folders for this instrument are stored on local disk.
- If the hostname is available on the network, then it will show as OK in the Available column.
- Add the instrument to MISO by clicking on the *Add* link.

### Editing a Sequencer Reference

- Editing a sequencer reference allows the administrator to change the instrument information such as the name and hostname.

## Working with MISO

In MISO, Project elements are the heart of the system, and comprise many components. All elements are interconnected, based on models designed from the submission schemas specified by the SRA.

# Project

A MISO project is a collection of studies, samples, and libraries. Each project has a *prefix* and an *ID*. The project ID becomes the suffix of the project's name, e.g. **PRO1**, **PRO2**. The project name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any project is **[http://your.miso.host/miso/project/<project\\_ID>](http://your.miso.host/miso/project/<project_ID>)**.

## Creating a project

- Click on the *Add Project* link at the top-right of the My Projects tab
- A ID and name will be auto-generated for the project upon saving.
- Alias - Important information about Project or Specific Project name,
- Description
- Set a progress value
- Add an Issue Tracker key, project can be created by relevant JIRA issue information.
- Set permissions to other user for access of the project, special group called "Internal" can be used here with a tick to allow all internal staff to access the project.
- Click *Save* to save the new project. The following elements can now be edited for the project.
  - Documentation files related to the project can be uploaded in the section
  - A Project Overview can now be added to the project, as well as Samples and Studies.

## Adding a Project Overview

- A Project Overview represents the portion of work to be done on a particular number of samples for a particular collaborator.
- Once an overview has been added, this portion of the project work can now be tracked via the alerting system, and will be reflected in the progress indicator.

## Samples

- Click on the *Samples* expanding link on the right hand side of the page
- Hover over the *Options* menu
- Select *Add Sample* to start adding samples following the documentation below

## Libraries

- Click on the *Samples* expanding link on the right hand side of the page
- Hover over the *Options* menu
- Select *Bulk Add Libraries* to start adding Libraries following the documentation below

## Library Dilutions

- Click on the *Libraries* expanding link on the right hand side of the page
- Hover over the *Options* menu
- Select *Bulk Add Library Dilutions* to start adding dilutions following the documentation below

## EmPCRs

- Click on the *Library Dilutions* expanding link on the right hand side of the page
- Hover over the *Options* menu
- Select *Add EmPCR* to start adding EmPCRs following the documentation below

## EmPCR Dilutions

- Click on the *EmPCRs* expanding link on the right hand side of the page
- Hover over the *Options* menu
- Select *Add EmPCR Dilution* to start adding EmPCR Dilutions following the documentation below

## Pools

- Click on the *Library Dilutions* or *EmPCR Dilutions* expanding links on the right hand side of the page
- Hover over the *Options* menu
- Select *Create Pools* to start adding creating Pools following the documentation below

## Studies

- Click on the *Studies* expanding link on the right hand side of the page
- Hover over the *Options* menu
- Select *Add Study* to start adding studies following the documentation below

## Runs

- Any runs associated with the current project, i.e. runs which have pools containing samples from this project, will be shown in this area.

## Sample

A MISO sample represents the physical material received upon which QCs and library preparations are carried out to prepare that sample for eventual sequencing. Each sample has a *prefix* and an *ID*. The sample ID becomes the suffix of the sample's name, e.g. **SAM1**, **SAM2**. The sample name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any sample is [http://your.miso.host/miso/sample/<sample\\_ID>](http://your.miso.host/miso/sample/<sample_ID>).

### Creating a single sample

- A sample is related to a specific Project and is tracked by a unique **alias** and **identification barcode**.
- Add in a sample description, outlining perhaps the origin and eventual preparation of this sample.
- Add a sample alias. Sample aliases must conform to the following specification:
  - <collaborator\_identifier>\_S<sample\_number\_relative\_to\_parent\_project>\_<material\_identifier>
  - where **collaborator\_identifier** might be a PI's initials
  - where **sample\_number\_relative\_to\_parent\_project** is a sample number supplied for the given set of samples within a project, e.g. 1 for the first sample, 2 for the second etc
  - where **material\_identifier** is usually the short species name, or some other sensible description
  - i.e. ABC\_S1\_H.sapiens
  - A sample alias cannot contain reserved characters or spaces.
- Specify a receipt date if this sample has already been received. Leave blank if this sample is to be received at a later date, i.e. when the sample delivery form has been returned to your centre.
- Specify a NCBI-valid species name for this sample. If enabled, MISO will check this species with the NCBI and ascribe the returned taxon ID to this sample.
- Specify the type of this sample.
- Lastly, specify whether this sample has passed QC. Please see the Sample QC section on how to add Sample QCs

### Bulk sample creation

- Bulk sample addition is identical to single creation, but in tabular form for ease of adding multiple samples.
- To start the process, click on *Add Row* at the top of the table. An empty row will be added.
- Add in the required fields as you would a normal single sample input.
- To copy this row, click the arrow icon in the *Copy* field of the table. You will see that by default, the **sample\_number\_relative\_to\_parent\_project** in the sample alias is incremented automatically. To turn this feature off, expand the *Table Options* section above the table, and deselect the *Increment Sample Aliases Automatically* option.
- Alternatively, you can specify a defined number of rows to add immediately by clicking the *Bulk Copy* button. A dialog will appear asking you how many rows to add. Choose a value, e.g. "96", and click *OK*. This will copy the last row in the table x times, incrementing the **sample\_number\_relative\_to\_parent\_project**.
- To quickly fill down a value in a field to other Samples, fill in the value for the chosen field in an upper row, select that row, and then click the *Fill Down* arrow button in a column header. All Samples below the chosen row will have the same value applied to that field.
- Click the *Save* button. You will be informed of any rows that did not save.

### Sample operations

Once a Sample has been saved, the following operations are available when visiting the Edit Sample page, via clicking on a sample's *Edit* link in any relevant listing table, e.g. from the List Samples page, or the Samples table in a relevant Project.

### Printing Sample Barcodes

- Once a Sample has been saved, a barcode will be generated for it. This will be present in the top right of the screen when editing a Sample.
- Hovering over the barcode will show a menu with the *Print* option.
- Clicking on this will show a dialog allowing the user to select which printer with which to print the barcode.

### Adding Notes

- Click on *Add Notes* under the Notes header.
- Add any free-text notes related to the Sample.

### Adding Sample QCs

- In the Edit Sample Page, Click on *Add Sample QC* under the Option menu.
- Add details
  - QCed By, QC Date, QC Method and Result

## Adding Libraries

- In the Edit Sample Page, Click on *Add Library* under Option of Libraries Pane.
- Select *Add Library* to start adding Libraries to start adding samples following the documentation below

## Receiving samples

- Visit the *Receive Samples* link in the Navigation Bar to access the form to receive samples.
- Either type, or beep in using a barcode reader, a sample barcode.
- If this sample has no previously set receipt date, it will be added to the list below the input box.
- You can repeat this process for as many samples as you need to receive.
- Once finished inputting, click *Save*.

## Library

A MISO sample represents the first step in constructing sequenceable material from an initial Sample. A Library is then diluted down to a Dilution, and put in a Pool.. Each library has a *prefix* and an *ID*. The library ID becomes the suffix of the library's name, e.g. **LIB1**, **LIB2**. The library name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any sample is [http://your.miso.host/miso/library/<library\\_ID>](http://your.miso.host/miso/library/<library_ID>).

## Adding Libraries

- A Library is related to a specific Sample and is tracked by a unique **alias** and **identification barcode**.
- Add in a library description, outlining perhaps the origin and eventual preparation of this library.
- The Library alias inherits from the parent Sample, but with some important changes:
  - A Library alias conforms to  
`<collaborator_identifier>_L<parent_sample_number>-<library_number_relative_to_parent_sample>_<material_identifier>`
  - Thus the first Library added to ABC\_S1\_H.sapiens would have an alias of ABC\_L1-1\_H.sapiens
  - The second would be ABC\_L1-2\_H.sapiens, and so on
- The creation date is automatically set to the current date.
- Check the Paired checkbox if the library is paired.
- Set the Platform that this Library is being prepared for, the Library type, Library section type, Library strategy type.
- If this Library is to be multiplexed with an adaptor barcode, pick the barcode from the dropdown list available for the selected platform.
- Check the QC Passed checkbox if the library has passed and QC stages.
- Permissions will be inherited from the parent Sample.

## Bulk Library creation

- Bulk Library addition is identical to single creation, but in tabular form for ease of adding multiple Libraries.
- You will see the list of Samples associated with the parent Project. These form the basis for Library creation.
- Add in the required fields as you would a normal single Library input.
- Select the samples you wish to use for Library creation by clicking in the *Select* column so that the row turns blue.
- **One** Library will be created for each Sample selected as such in the table.
- To quickly fill down a value in a field to other Libraries, fill in the value for the chosen field in an upper row, select that row, and then click the *Fill Down* arrow button in a column header. All Libraries below the chosen row will have the same value applied to that field.
- Click the *Save* button. You will informed of any rows that did not save.

## Library Operations

Once a Library has been saved, the following operations are available when visiting the Edit Library page, via clicking on a Library's *Edit* link in any relevant listing table, e.g. from the List Libraries page, or the Libraries table in a relevant Project.

## Adding Notes

- Click on *Add Notes* under the Notes header.
- Add any free-text notes related to the Library.

## Adding Library QCs

- Click on *Add Library QC* under Option menu.
- Add details
  - QC Date, QC Method and Result.
  - Insert Size in base pairs.
  - Click on *Add* to save.

## Add Library Dilutions

- Click on *Add Library Dilution* under Option of Library Dilutions Pane.

- Add details
  - Creation Date and Result
  - The ID barcode for the dilution will be generated on save
  - Click on *Add* to save.

## Constructing Pools

- Once a Library Dilution has been saved, you can use it to make up a pool
- Click on the *Construct new Pool using this dilution...* link to go to the Pool Wizard.
- Please see the Pool Wizard documentation below.

## Pool

A MISO Pool contains one or more Dilutions that are to be placed, as part of an Experiment, in a sequencer instrument Run partition (lane/chamber). Pools with more than one Dilution are said to be multiplexed. Each pool has a *prefix* based on the Platform Type and an *ID*. For Illumina pools, the prefix will be **IPO**, for 454 **LPO**, for SOLiD **SPO**. The pool ID becomes the suffix of the pool name, e.g. **IPO1**, **LPO270**. The pool name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any sample is [http://your.miso.host/miso/pool/<platform\\_type\\_in\\_lower\\_case>/<pool\\_ID>](http://your.miso.host/miso/pool/<platform_type_in_lower_case>/<pool_ID>).

## Creating Pools

- A Pool is related to a specific Experiment and is tracked by a unique **alias** and **identification barcode**.
- On creation, the Pool ID and name will be auto-generated.
- Add in a Pool alias, concentration and creation date.
- Check the Ready to Run checkbox if this Pool is ready to be run on a sequencer, i.e. no more dilutions need to be added.
- If an Experiment has already been created for this Pool, use the *Search Experiments* field to find valid Experiments based on the name, description or related project.
  - Once found, click on the Experiment to link it to the Pool.
  - A pool can be linked to multiple experiments to denote a multiplexed pool comprising dilutions from different projects.
- If an Experiment has **NOT** been created for this Pool, you can automatically create one simply by adding the pool to a Run partition, and selecting the related study/ies. An Experiment will then be auto-generated for you for each Study.
  - **NB** This process requires that at least one Study is present for each of the projects related to this Pool via the prepared Samples.
- Add Dilutions to the Pool by:
  - Using the *Search Dilution* field to find valid Dilutions based on the name, description or related project
  - Using the *Select Dilutions by Barcodes* to beep in multiple ID barcodes of dilutions to be added. Barcodes should be on separate lines. Once all dilutions have been beeped in, click *Select* to add them.
  - Using the *Select Dilutions by Barcode File* chooser to upload a file with dilution barcodes in. Barcodes should be on separate lines within the file.
  - **NB** Dilutions can only be added to the same Pool if the parent Library objects have different Tag Barcodes.
- Both Experiments and Dilutions can be removed from the Pool at any time by clicking on the cross icon on the right of the Experiment or Dilution, and confirming the resulting dialog box.

## Pool Operations

Once a Pool has been saved, the following operations are available when visiting the Edit Pool page, via clicking on a Pool's *Edit* link in any relevant listing table, e.g. from the List Pools page, or the Pools table in a relevant Project.

## Study

A MISO study represents more fine-grained information about the sequencing Project. Each study has a *prefix* and an *ID*. The study ID becomes the suffix of the study name, e.g. **STU1**, **STU2**. The study name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any sample is [http://your.miso.host/miso/study/<study\\_ID>](http://your.miso.host/miso/study/<study_ID>).

## Creating a study

- A study is related to a specific Project and is tracked by a unique **alias**.
- Add in a Study description, outlining perhaps the origin and purpose of this Study.
- Add a Study alias.
  - A Study alias cannot contain reserved characters or spaces.
- Specify the type of this Study.

## Edit Study

Once a Study has been saved, the following operations are available when visiting the Edit Study page, via clicking on a Study's *Edit* link in any relevant listing table, e.g. from the List Studies page, or the Studies table in a relevant Project.

## Adding Experiments

- Studies can contain any number of sequencing Experiments.
- Select *Add Experiment* to start adding Experiments following the documentation below.

## Experiment

A MISO Experiment represents design information about the sequencing experiment. Each Experiment has a *prefix* and an *ID*. The Experiment ID becomes the suffix of the Experiment's name, e.g. **EXP1**, **EXP2**. The Experiment name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any Experiment is [http://your.miso.host/miso/experiment/<experiment\\_ID>](http://your.miso.host/miso/experiment/<experiment_ID>).

Experiments are associated with Pools that are placed on Run partitions, which contain the actual sequencing results.

## Automatic Experiment Creation

- Whilst MISO supports creating Experiments manually (see below), it is often easier to allow MISO to generate and link an Experiment to a Pool for you.
- From the *Edit Run* page, you are able to select Pools to be placed on Run partitions (lanes/chambers).
- When assigning a Pool to a partition, the following rules apply for Experiment creation:
  - If the Pool has no Experiment already linked
  - If there is at least one Study available on each of the Projects represented by the Pool's dilutions
- If these rules are valid, drop boxes will appear for each represented Project, letting you choose the Study on which to base the automatic Experiment creation.
- Upon clicking *Select* for each Study, one Experiment will be automatically created and linked to the Pool.

## Add Experiment Manually

- An Experiment is related to a specific Study and is tracked by a unique **alias**.
  - An Experiment alias cannot contain reserved characters or spaces.
- Add in an Experiment title, alias, and description.
- Specify the platform to which this Experiment is related.
- Selecting the Platform will populate a list of available Pools to link to this Experiment.
- Double click to add a chosen Pool to this Experiment.
- Double clicking again on the chosen Pool area will remove it.
- Permissions for each Experiment are inherited from the parent Study

## Partition Containers

Partition Containers represent a collection of Partitions (abstractions of lanes, chambers, etc) that are described by a parent Run object. Containers can be populated prior to a Run being entered manually or being imported by the Notification System. This means Containers can be barcoded and entered into MISO with their Pools populated ahead-of-time so that when Runs are added manually or automatically, the Containers can either be linked via a quick lookup or automatically, respectively.

## Adding Partition Containers (outside a Run)

- Independent Container creation is accessed from the *Create New Partition Container* link in the navigation area
- Select a Platform for the Container from the *Platform* radio buttons
- Select an available Sequencer Reference from the *Sequencer* dropdown
- Select the number of Containers you want to add for the Platform (454 is the only multiple flowcell vendor. HiSeq's are effectively "single flowcell in twos", i.e. two run folders are created independently even though two flowcells can be run concurrently on the machine itself).
- Enter in an ID for this Container. You can validate that this barcode hasn't been used before by using the *Lookup* button.
- A location and validation barcode can also be entered, but are not compulsory.
- Double click on an empty partition and either beep in or type the barcode of the pool you wish to add to this Container
- Click on the *Save* button to save.

## Adding Partition Containers (within a Run)

- Containers can also be created and edited from within a Run itself and the process is very similar to the above.
- Select the number of Containers you want to add for the Platform (454 is the only multiple flowcell vendor. HiSeq's are effectively "single flowcell in twos", i.e. two run folders are created independently even though two flowcells can be run concurrently on the machine itself).
- Enter in an ID for this Container. You can validate that this barcode hasn't been used before by using the *Lookup* button.
- A location and validation barcode can also be entered, but are not compulsory.
- Pools are added to Container partitions from the Pool list on the right hand side of the page by double clicking on a Pool
- **NB** the chosen Pools will be added **sequentially** to the Container, as by this point, MISO will assess that as the run is already started, it should be fully populated.
- If you wish to retrospectively semi-populate a Container, use the method in the previous section ("outside a Run").

## Run

A MISO Run contains the sequencing results from sequencing experiments. Each Run comprises one or more Partition Containers, that in turn comprise Partitions (abstractions of lanes, chambers, etc) which hold a Pool. The Run ID becomes the suffix of the Run name, e.g. **RUN1**, **RUN2**. The Run name is programatically constructed to be unique and as such is unchangeable by any user, including administrators. The URL address for viewing any Run is [http://your.miso.host/miso/run/<run\\_ID>](http://your.miso.host/miso/run/<run_ID>).

### Automatic Run Notification

MISO is bundled with a separate Java application that can be run on a machine with access to sequencer run folders. This Notification module allows MISO to be informed about new Runs without any involvement from a user, making it far simpler to manage Runs within MISO. More information on this system can be found in the relevant Notification System section.

### Adding a Run Manually

- A Run is related to Projects via the pooled Samples placed on its partitions and is tracked by a unique **alias**.
- An ID and name will be auto-generated for the Run upon saving.
- Select a Platform Type.
- Select a Sequencer Reference
  - Once a Sequencer Reference has been selected, you will be able to select and populate the Container(s). Please follow the information in the Partition Containers section above.
- A Run alias, which should be the run name specified by the sequencer system, and usually relates to a run folder of the same name on the underlying filesystem of use.
- A Run description, giving a brief overview of what the Run represents.
- The Run path, which should be the run folder exposed by the sequencer itself or an intermediary storage filesystem or cluster.
- Check the Paired End checkbox if the Run is paired end.
- Select the current Run status.
- Click on Save to save the Run.

### Run Operations

Once a Run has been saved, either manually or via the Notification System, the following operations are available when visiting the Edit Run page, via clicking on a Run's *Edit* link in any relevant listing table, e.g. from the List Runs page, or the Runs table in a relevant Project.

- Edit alias and description
- Choose if the Run is paired end
- Setting the Status of the Run manually
  - This is also helpful if the notification system fails to accurately record the status of a Run (some platforms don't expose this information in a coherent fashion).

### Editing Partition Containers

- See the Partition Containers section above.

### Adding Notes

- Click on *Add Notes* under the Notes header.
- Add any free-text notes related to the Run.

### Adding Run QCs

- Click on *Add Run QC* under the Option menu.
- Add details
  - The QC User that carried out the QC.
  - The QC date.
  - The QC Method relates to a conceptual or practical QC steps undertaken by one of two "actioners", i.e. SeqOps (*sequencing operations*) or SeqInfo (*sequence informatics*). The former is carried out after the physical run has completed and assessed for quality. The latter is carried out following a primary analysis stage to assess quality.
  - The Process Selection allows the QC to be related to a given subset of partitions, e.g. if lanes 1-4 were OK and should be processed, but lanes 5-8 were not, boxes 1-4 would be clicked to reflect this.
  - Any other free-text information should be added in the Info field.
  - Click on *Add* to save.

## Reporting

Reporting in MISO is a very flexible feature, with reports that can be generated based on Projects, Samples, Libraries and Runs. All reports can

be filtered by keywords, date range and relevant types.

## Generating Reports

1. Reports based on different elements can be generated from separate tabs for each, i.e. Project report, Sample Report, Libraries Report and Runs Report
2. Each type of report got different filtering options
  - a. All type of reports can be filtered by using keyword for name, alias and description, and dates
  - b. Project reports got extra filter option for Project Progress, i.e. Active, Inactive, Pending, Cancelled, Proposed, completed, Approved, Unknown.
  - c. Samples reports can be filtered based on Samples Type and Sample QC
  - d. Libraries reports filters are Sequencer Platform type and Library QC
  - e. Runs can be filtered based on Run Status, and Sequencers Platform type
3. Press the *Search* button, will give you a list of objects based on filters,
4. *Reset* button clears all the filters and reset the list,
5. Select the relevant objects or select *All* check box and press *Generate Report*, the report will then be produced.

## Report output

Includes

- Statistics table based on the state of Projects, Sample, Library, Runs
  - with doughnut chart(s)
- Tabular representation of all elements display information for each element separately.
  - This can be search based on keywords

## Sequencer Resource Calendar

- The Resource Calendar shows Run information for each sequencer available in MISO.
- Results can be filtered:
  - Based on Instrument
    - Single Instrument or all
    - For single instrument view in multi-flowcell instruments, e.g. HiSeq2000, both cells are shown on separate rows.
  - Based on time
    - Set time periods, i.e. Current Year, Last Year, Last Week, Last Month, Last Three Months, Last Six Months
    - In Custom, start and end dates can be selected independently

## Submissions

MISO has an automated submission feature, enabling you to quickly and easily submit your sequence meta data to public repositories such as the EBI. MISO will automatically generate the required XML documents based on the information already stored in the system, and submit these to the repository. MISO can also be configured to initiate upload of your actual sequence data.

## Preparing new Submissions

- To prepare a new Submission, go to the *My Account* tab, and choose *Prepare new Submission*.
- Enter a title, alias and description for your submission in the appropriate boxes.
- Choose the action for the submission from the options: Add, Modify, suppress, Hold, Release, Close, Protect or Validate.
- You can now choose which elements you want to submit:
  - From the list of Projects, clicking on any of these will reveal a drop-down list of submittable elements.
  - Check the boxes next to the elements you wish to submit, and click *Save*.

## Previewing, Validating and Submitting Metadata

- Once you have saved your Submission, you will see links to either Preview, Validate or Submit your submission metadata.
  - **Preview** will show you a condensed version of the metadata in a new browser window.
  - **Validate** will generate the submission XMLs and send them to the EBI validation service. Any problems with the XMLs will be displayed in the browser window - errors in Red and information in Yellow. The actual metadata will be displayed below these warnings.
  - **Submit Metadata** will start the metadata submission process. This will also be validated, and will fail, displaying error messages if the XML is invalid.

## Submitting Sequence Data

- Once you have saved your Submission, you will see a link to submit your sequence data.
- MISO has a number of mechanisms to find data files based on the platform and submittable elements selected.
- These mechanisms should be suitable for most platforms outputting data to standard locations. If not, please see the Developer Manual section on writing custom FilePathGenerators.
- MISO also has mechanisms for uploading data to submission endpoints, e.g. the EBI SRA. MISO ships with a built-in FTP upload feature that will connect to an FTP server and upload datafiles. We are currently working on implementing other, faster upload methods such as ASPIRE.

## List Submissions

- Choose *List Submissions* from the admin area to see a list of all submissions that have been created so far, along with their status. You can edit any of these by clicking the *Edit* link.

## Notification Server

The Notification Server module for MISO collects up sequencer run information from a given set of datapaths (directories on a filesystem) and sends that information to MISO in a simple JSON format, which consumes it and automatically creates and updates Run objects as necessary.

## Installing

- Grab the notification-server tarball from the MISO repos site: <https://repos.tgac.bbsrc.ac.uk/miso/notification-server-0.1.6RC.tar.gz>
- Unpack it on a machine that has access to your sequencer's run folders.
  - For example, at TGAC we run it on a specialised virtual machine that has the output folders mounted as a network share. You could just as easily run this module on a cluster head node, or indeed an actual sequencer control PC.
- Copy the **example-notification.properties** to a new file called **notification.properties**.
- Edit the information in the properties file to match your desired configuration. See *Configuring* below.

## Configuring

- The supported platforms are: **illumina**, **solid**, **Is454**, **pacbio**. Each subsequent property is demarcated by this platform identifier.
  - **wiretap.enabled**:false
    - Enabling this global option will give you very verbose output about what is getting packaged up and sent to the MISO server. Only turn this on if you like lots of eye-burning developer output.
  - **illumina.dataPaths**:/path/to/your/illumina/run/folders
    - Set this property to one or more (comma separated) paths on the filesystem where your Illumina run folders are stored.
  - **illumina.http.statusEndpointURLs**:<http://your.miso.server/miso/consumer/illumina/run/status>
    - Set this property to one or more (comma separated) URLs that represent endpoints to send the run information to. This is usually a MISO install, but could be any JSON endpoint you choose. Please see the Developer Documentation for the JSON schema if you want to develop your own endpoints.
  - **illumina.scanRate**:600000
    - Sets the period in seconds to poll the dataPaths above for new folders
  - **solid.dataPaths**:/path/to/your/solid/run/folders
    - As above.
  - **solid.http.statusEndpointURLs**:<http://your.miso.server/miso/consumer/solid/run/status>
    - As above.
  - **solid.wsdl.url.s0000**:<http://your.solid.machine1:8080/sets/webservice/solid?wsdl>
    - Some SOLiD platforms expose web services that can be queried for information. Enter the wsdl path here. The part of property key after the 'url', i.e. **solid.wsdl.url.s0000** needs to EXACTLY MATCH the sequencer reference name held in MISO for this machine. There can be more than one of these properties, each with a unique sequencer reference identifier.
  - **solid.scanRate**:600000
    - As above.
  - **Is454.dataPaths**:/path/to/your/454/run/folders
    - As above.
  - **Is454.http.statusEndpointURLs**:<http://your.miso.server/miso/consumer/Is454/run/status>
    - As above.
  - **Is454.scanRate**:600000
    - As above.
  - **pacbio.dataPaths**:/path/to/your/pacbio/run/folders
    - As above.
  - **pacbio.http.statusEndpointURLs**:<http://your.miso.server/miso/consumer/pacbio/run/status>
    - As above.
  - **pacbio.ws.url.12345**:<http://your.pacbio.machine1/>
    - As with the SOLiD web services url, this property represents the PacBio RS REST API endpoint. The part of property key after the 'url', i.e. **pacbio.ws.url.12345** needs to EXACTLY MATCH the sequencer reference name held in MISO for

- this machine. There can be more than one of these properties, each with a unique sequencer reference identifier.
- **pacbio.scanRate:600000**
  - As above.

## Running

Simply run the **run\_notification\_server.sh** shell script. We usually do this in a **screen** (\*nix) so you can detach the process and come back to it later on. If you are on Windows, knock up a quick batch script that does the same thing!

## Analysis Server

The Analysis Server module enables MISO to perform workflows as simple pipeline tasks, comprising one or more processes, and fire these off to an LSF machine. This code is heavily based on the ArrayExpress Conan architecture, developed at the EBI by Tony Burdett.

## Installing

- Grab the analysis-server tarball from the MISO repos site: <https://repos.tgac.bbsrc.ac.uk/miso/analysis-server-0.1.6RC.tar.gz>
- Unpack it on a machine that has the ability to run jobs via LSF (bsub).
  - For example, at TGAC we run this on a machine that can call our LSF master.
- Copy the **example-analysis.properties** to a new file called **analysis.properties**.
- Set up a new MySQL instance on a server (or use an existing one) and install the Conan database as described in the MISO install [README](#)
- Edit the information in the properties file to match your desired configuration. See *Configuring* below.

## Configuring

- **analysis.server.port:7898**
  - The port that the server will listen for queries.
- **analysis.client.host:localhost**
  - The client socket hostname that will send out responses (usually 'localhost')
- **analysis.client.port:7898**
  - The socket port of the client that sends out responses (usually the same as the server port)
- **analysis.db.url:jdbc:mysql://your.conan.db.server:3306/conan**
  - Conan writes task and process tracking information to a SQL database. Supply the connection string to that database here
- **analysis.db.username:conan**
  - The Conan db username
- **analysis.db.password:conan**
  - The Conan db password
- **analysis.server.maxMessageSize:20480**
  - The maximum message size that can be received by the server
- **analysis.client.maxMessageSize:20480**
  - The maximum message size that can be sent out by the client
- **analysis.submission.parallelJobLimit:30**
  - The maximum number of parallel jobs that can be processed.
- **analysis.submission.coolingOffPeriod:120**
  - The length of time in seconds between submitting a job submission task request and the job being sent to LSF.

## Running

Simply run the **run\_analysis\_server.sh** shell script. We usually do this in a **screen** (\*nix) so you can detach the process and come back to it later on. If you are on Windows, knock up a quick batch script that does the same thing!

## Troubleshooting

[Go to start of metadata](#)

- The current MISO is a beta version so as such doesn't guarantee bug free usage.
- If a bug has been found it can be reported to the developers via the MISO JIRA issue system.
- If you are a developer and have found a bug and have fixed it, please submit a patch via the MISO JIRA system.

<https://repos.tgac.bbsrc.ac.uk/miso/notification-server-0.1.6RC.tar.gz>